DOCUMENT RESUME

ED 232 145 CS 007 234

AUTHOR Christine, Charles T.; And Others
TITLE The Between Teacher Reliability of the Ekwall Reading Inventory and the Classroom Reading Inventory.
PUB DATE Oct 82
NOTE 15p.; Paper presented at the Annual Meeting of the Keystone State Reading Association (15th, White Haven, PA, October 13-15, 1982).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; Elementary Secondary Education; *Informal Reading Inventories; Reading Instruction; *Reading Research; Reading Teachers; Reading Tests; *Test Reliability
IDENTIFIERS *Classroom Reading Inventory; *Ekwall Reading Inventory

ABSTRACT

Using a test-retest research design, a study examined the reliability of the Classroom Reading Inventory (CRI) and the Ekwall Reading Inventory (ERI). Independent variables of test administrator to subject, test administrator to test, subject to test, and test order were randomized. Subjects included 31 children aged 7 through 12 years. The four teachers who served as examiners were all graduates of master's degree programs in developmental and remedial reading instruction. The test was a "live" administration of one of the reading inventories by one teacher; the testing session was tape recorded. The retest was a second teacher scoring the audio tape recording of the test. The dependent variable was the agreement (or lack of agreement) in identifying a single reading instructional level between test and retest. Results showed that in 14 of 16 trials (85%) there was agreement between teachers using the ERI. In 5 of the 16 trials there was perfect agreement between teachers on an instructional reading level. None of the test-retest trials showed a teacher-teacher instructional level difference of more than one grade level. Trials of the CRI showed teacher agreement on instructional level in 10 of 16 trials (67%). Only one test-retest instance showed a between-teacher instructional level disagreement of more than one grade level. These results provide a strong indication that the CRI and ERI produce reliable estimates of a pupil's reading instructional level. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED232145

The Between Teacher Reliability of the Ekwall Reading Inventory and the Classroom Reading Inventory

Charles T. Christine
Lawrence A. Anderson
Edythe Bleznak
Jane B. Levine
Phyllis Lewy

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Charles T. Christine

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS 007 234

The Between Teacher Reliability of the Ekwall Reading Inventory and the Classroom Reading Inventory

The informal reading inventory (IRI) is a criterion referenced test of a pupil's ability to read and comprehend passages written at a graded level of difficulty (Johnson and Kress, 1965). Two types of IRI's are in general use: (1) IRI's keyed to specific reading practice materials such as a basal reader set; and (2) IRI's designed to identify a pupil's instructional reading level for a broad range of reading materials.

Information on the pupil's instructional reading level is useful to the reading teacher. Such information is needed to select practice material sufficiently difficult to challenge the pupil, but not so difficult as to frustrate the pupil's efforts to read and comprehend (Johns, 1977). The IRI is used to gather information to support instructional reading level decisions for specific pupils, and it is often used to determine a "starting point" of graded practice materials for pupils new to a reading program.

There is little question that the IRI content is suited to determining pupil instructional reading level. The IRI text is referenced to the graded reading materials that the pupil will use. In this sense it is a criterion referenced test. However, there is some question of how consistently a given IRI places a pupil's reading instructional level. This is a question of IRI reliability.

McRae analyzed the reliability of the Houghton-Mifflin Pupil Placement Tests, a basal-text oriented IRI. His study was an analysis of 36 administrations of the Pupil Placement Tests using a split-half reliability study model. McRae reported a Spearman-Brown coefficient of correlation of .90 for the Pupil Placement Tests (McRae, 1974).

West studied the reliability of the West Informal Reading Evaluation, an IRI designed for adult readers. West tested 100 adult subjects using a test-retest model of reliability study. West reported a test-retest coefficient of correlation of .85 for these trials (West, 1977).

Ekwall subjected the Ekwall Reading Inventory (ERI) to a reliability study using the alternate forms reliability model. Ekwall reported alternate forms reliability coefficients of .82 for one study and .79 for a second study of the ERI (Ekwall, 1979).

Christine and Bleznak studied the reliability of the ERI and the Classroom Reading Inventory (CRI), two commercially published IRI's. The researchers completed a reliability study based upon the test-retest model. However, instead of computing coefficients of correlation, the researchers simply reported gross data. For this study, Christine and Bleznak found that in twenty-seven of thirty-two cases (84% of the trails), a second administration of the ERI or CRI placed a child within one grade level of a first testing (Christine and Bleznak, 1982).

McCracken demonstrated the reliability of the Standard Reading Inventory, a comercially distributed IRI. He computed Pearson Product Moment coefficients of correlation from a test-retest study of two teachers administering the Standard Reading Inventory to sixty children. McCracken reported a test-retest correlation of .91 for determinations of instructional reading levels from the SRI (McCracken, 1966).

All of the above studies, save McCracken's, involved one teacher testing one child either one time or two times with the same test as a way of determining IRI reliability. A more challenging study of test reliability is to repeat a test-retest sequence for one child with different teachers giving the test and retest. Use of this model makes possible an estimate of the degree to which different teachers agree or disagree as they use one type of IRI to determine instructional reading levels for their pupils. Within the above test-retest model, a wide disagreement between teachers on instructional reading level would call into question the usefulness of the IRI as a factor in practice materials selection. Conversely, close agreement between teachers on instructional reading level would suggest that the IRI gives reliable results between teachers and that attempts to match pupil reading ability with the difficulty level of practice materials are reliable.

This matching of different teachers' IRI administrations is what McCracken did in his study of SRI reliability. However,

McCracken used a Pearson-Product Moment comparison which has since been called into question as a proper method of reliability study. According to Nunnally, the use of the coefficient of correlation produces a distorted estimate of test-test reliability (Nunnally, 1971).

The trend of the research reported above suggested a new research design for the study of IRI reliability. There is a need for an IRI reliability study model that: (1) involves a random assignment of teachers to test-retest treatments; (2) uses a statistic other than the coefficient of correlation to represent the extent of test-retest agreement; and (3) allows for the test-retest comparison of one IRI administered to one subject by two teachers.

The Research Design: The design for the study reported herein was structured to isolate IRI reliability as the dependent variable of the study. Independent variables of test administrator to subject, test administrator to test, subject to test and test order were randomized.

Subjects were thirty-one children ages seven through twelve years of age. The subjects were randomly selected from a pool of volunteers attending two summer camps. None of the volunteers was mentally retarded. Instructional reading levels for the volunteer subjects ranged from pre-primer through adult.

The four teachers serving as examiners for the study are all graduates of masters degree programs in developmental and remedial

reading instruction. Their graduate programs included at least one academic quarter of supervised clinical practice. All four teachers are certified as specialist reading teachers in the Commonwealth of Pennsylvania. Three of the teachers are female and one is a male.

The research design was a test-retest model. The test was a "live" administration of an IRI by one teacher. This administration was tape recorded. The retest was a second teacher scoring the audio tape recording of the test. The dependent variable was the agreement (or lack of agreement) in identifying a single reading instructional level between test and retest. The researchers decided that test-retest instructional level agreement within two grade levels (plus or minus one level) would constitute a reliable result. This plus-or-minus one grade level decision was a practical clinical judgment on the part of the researchers. No guidance from the literature is available on this matter.

Informal reading inventories tested for reliability in this study were the CRI and the ERI. Each subject received one administration of one or the other test. This first administration was audio tape recorded and counted as the "test". A second scoring of the audio tape was completed by a different teacher, and this second scoring counted as the "retest".

A complex random assignment control table was constructed prior to testing. This table insured:

1. an equal, but random distribution of pupils among ERI and CRI test administrations;

2. an equal, but random distribution of teachers to pupils and to test-types;
3. a randomized order of ERI or CRI test administrations by the four teachers.

The final step prior to testing was the random assignment of teachers and pupils to the previously lettered and numbered "test slots" in the control table of assignments. Thus, each pupil had an equal opportunity to be assigned to either a CRI or an ERI and an equal opportunity to be tested by any one of the four teachers. Each test had an equal opportunity to be given to any pupil by any teacher.

Results of the Research: The between teacher test-retest reliability trials of the Ekwall Reading Inventory are summarized in Table I. In fourteen of sixteen trials (85%) there was agreement between teachers using the ERI. In five of the sixteen trials there was perfect agreement between teachers on an instructional reading level. None of the test-retest trials showed a teacher-teacher instructional level difference of more than one grade level. Rather, in two of sixteen cases neither teacher was able to definitively isolate any instructional level at all for either of the children. If this inability to determine an instructional level is counted as teacher agreement, then the teacher-teacher reliability of the ERI increases to sixteen of sixteen trials (100%).

TABLE I

The Between Teacher Reliability of the Ekwall Reading Inventory

Instructional Level Test Results	Number of Test-Retest Occurances
Perfect Agreement	5
Within One Level	9
Within Two Levels	0
No Level Isolated	2
Total N	16

Table II summarizes the between teacher test-retest reliability experience for the Classroom Reading Inventory (CRI). Trials of the CRI showed teacher agreement on instructional reading level in ten of sixteen trials (67%). If instances of failure to isolate an instructional reading level are counted as agreement, then test-retest agreement from these trials increases to fourteen of fifteen trials (93%). Only one test-retest instance showed a between teacher instructional level disagreement of more than one grade level.

Discussion: This research is in agreement with previous studies of IRI reliability. This study differs from previous IRI reliability studies in improved independent variable control. In addition, Nunnally's criticisms of the use of coefficients of correlation for test-retest studies were considered, and results from this study were reported as simple percentages (Nunnally, 1967). Nevertheless, the ERI and CRI met the criterion of clinical utility decided upon by the researchers.

The results of this study provide a strong indication that the CRI and ERI produce reliable estimates of pupil reading instructional level. Replications of the study are needed to further establish the reliabilities of the CRI and ERI.

Six children in this study defied precise reading instructional level placement from a single IRI administration. The phenomenon of

TABLE II

The Between Teacher Reliability of the

Classroom Reading Inventory

Instructional Level Test Results	Number of Test-Retest Occurances
Perfect Agreement	7
Within One Level	3
Within Two Levels	1
No Level Isolated	4
Total N	15

of a small group of children defying placement was also observed in a previous test-retest reliability study (Christine and Bleznak, 1982). For purposes of clinical practice it should be recognized that the determination of an initial instructional reading level cannot be obtained from a single IRI administration for some children (Jongsma and Jongsma, 1981). For future research, projects need to be designed to investigate whether or not the concept of a single instructional reading level is invalid for some children.

Summary: This is a report of a study of the between teacher test-test reliability of the Ekwall Reading Inventory and the Classroom Reading Inventory. This study supports a test-retest agreement within plus or minus one grade level in 88% of the observed instances for the ERI and in 67% of the instances for the CRI. If instances of teacher failure to isolate any instructional reading level are considered as agreement, the ERI test-retest result increases to 100% and the CRI result to 93% agreement. It is reasonable to count instances of no level isolation as agreement, thus the higher reliability instances obtain for these trials.

This study suggests a satisfactory reliability for practical clinical applications of the CRI and ERI.

REFERENCES

Botel, Morton. Botel Reading Inventory Administration Manual. New York: Follett Publishing, 1978.

Braswell, Charles N. "A Comparison of Two Methods of Administering An Informal Reading Inventory." Doctoral dissertation, University of South Carolina, 1976.

Christine, Charles and Edythe Bleznak. "The Teacher-Teacher Reliability of the ERI and CRI." mimeo, The Montgomery County Intermediate Unit, Erdenheim, PA 19118, 1982.

Ekwall, Eldon B. Ekwall Reading Inventory. Boston: Allyn-Bacon, 1979.

Ekwall, Eldon B. Ekwall Reading Inventory Manual of Directions. Boston: Allyn-Bacon, 1979.

Farr, Roger. Reading: What Can Be Measured? Newark, Delaware: International Reading Association, 1969.

Hardin, Veralee B. and W.B. Ames. "A Comparison of the Results of Two Oral Reading Tests." The Reading Teacher. vol. 22 (January 1969), pp. 329-334.

Johns, Jerry et al. Assessing Reading Behavior. Newark, Delaware: The International Reading Association, 1977.

Johns, Jerry L. "Matching Students With Books." Contemporary Education. vol. 48 (Spring, 1977) pp. 133-36.

Johnson, Marjorie and Roy A. Kress. Individual Reading Inventories. Newark, Delaware: International Reading Association, 1965.

Jongsma, Kathleen and Eugene Jongsma. "Test Review: Commercial Informal Reading Inventories." The Reading Teacher. vol. 34 (March, 1981) 697-705.

Kelly, Dean. "Using an Informal Reading Inventory to Place Children In Instructional Materials." in Reading Difficulties: Diagnosis, Correction and Remediation, edited by William K. Durr. Newark, Delaware: International Reading Association, 1970. pp. 111-19.

Kender, Joseph. "An Analysis of Factors Associated With Informal Reading Tests At The Eighth Grade Level." April, 1968. Microfiche Ed. 019206.

Lowell, Robert E. "Problems in Identifying Reading Levels With Informal Reading Inventories." in Reading Difficulties: Diagnosis, Correction and Remediation. edited by William K. Durr. Newark, Delaware: International Reading Association, 1970. pp. 120-26.

McRae, L.G. "A Measure of the Reliability of the Informal Reading Inventory." Doctoral dissertation, University of Northern Colorado, 1974.

Memory, David, Glen Powell and Byron Callaway. "A Study of the Characteristics of the Woodcock Reading Mastery Tests." Reading Improvement, vol. 17 (Spring, 1980) pp. 48-52.

Nunnally, Jum C. Psychometric Theory. New York:McGraw-Hill, 1967. pp. 193-220.

Pikulski, John. "A Critical Review: Informal Reading Inventories." The Reading Teacher, vol. 28 (Nov. 1974) pp. 617-21.

Powell, William R. "Validity of the IRI Reading Levels." Elementary English, vol 48 (October 1971) pp. 637-42.

Silvaroli, Nicholas. Classroom Reading Inventory. Dubuque, Iowa: William C. Brown, 1976.

Silvaroli, Nicholas J. Classroom Reading Inventory Manual Of Directions. Dubuque, Iowa: William C. Brown, 1976.

Smith, William E. and Michael D. Beck. "Determining Instructional Reading Level With the 1978 Metropolitan Achievement Tests." The Reading Teacher, vol. 34 (December 1980) pp. 313-19.

West, Ellen. "The Use of an Unobtrusive Screening Device to Approximate Reading Levels of Adults." Paper presented to the Annual Meeting of the National Reading Conference, New Orleans, La. December 1-3, 1977.

Authors: Charles T. Christine, Ph.D., Montgomery County Intermediate Unit, Montgomery Avenue and Paper Mill Road, Erdenheim, PA. 19118

Larry Anderson, B.S., M.S., Montgomery County Intermediate Unit

Edythe Bleznak, B.S., M.S., Montgomery County Interemdiate Unit

Phyllis Lewy, B.S., M.S., Montgomery County Intermediate Unit